DIVISION OF THE HUMANITIES AND SOCIAL SCIENCES

CALIFORNIA INSTITUTE OF TECHNOLOGY

PASADENA, CALIFORNIA 91125

COMMON KNOWLEDGE AND CONSENSUS WITH AGGREGATE STATISTICS*

Richard D. McKelvey

R. Talbot Page

Published as "Common Knowledge, Consensus, and Aggregate Information," Econometrica 54 (1986):109-127.

SOCIAL SCIENCE WORKING PAPER 493

October 1983

Revised May 1984

Abstract

We prove that if n individuals start with the same prior over a probability space, and then each observe private information, that for a class of admissible statistics, if a statistic of their posterior probabilities of an event becomes common knowledge, then everyone's posterior probabilites must be the same. The class of admissible statistics includes any statistic which is an invertible function of a stochastically monotone function. We also prove that if information partitions are finite, an iterative procedure of public announcement of the statistic—where the statistic is publicly announced and then individuals recompute posterior probabilities based on their previous information plus the announced value of the statistic—converges in a finite number of steps to the common knowledge situation described above. The result has applications to Delphi type processes for probability assessment, and to economic models in which private information becomes incorporated into an aggregate, publicly observed statistic such as a price or quantity in a market.

COMMON KNOWLEDGE, CONSENSUS AND AGGREGATE INFORMATION

Richard D. McKelvey
Talbot Page

California Institute of Technology

(May 14, 1984)

1. Introduction

Generally speaking, information held privately by an individual is useless to him until he acts upon it. When several individuals act on their private information, some of the information may become incorporated into a publicly observable statistic, such as a price. Individuals may then make inferences from the public information to augment their original private information.

In this paper we use the concept of common knowledge to study the aggregation of private information into a public statistic, and the "redistribution" of information that occurs as individuals make inferences from the aggregate, public information. We derive conditions under which information becomes so well redistributed that there is consensus (but not necessarily full revelation). The analysis applies to a wide variety of settings, including oligopoly pricing, parimutuel betting, a Delphi process, and some models of rational expectations. We discuss several examples where our regularity condition, along with the background assumption of common knowledge of initial structures, is satisfied.

The analysis suggests that there may be a tendency in markets, and other institutions that aggregate information into public statistics, toward consensus in beliefs. In rough terms this tendency may stand in parallel to the tendency in markets toward consensus in preferences (at the margin) as individuals equate their marginal rates of substitution. However, as we shall see, the conditions leading to consensus in beliefs are more complicated.

We consider the case when n individuals begin with the same prior on a probability space, and then each obtain private information. They each take an action based on this information, and an aggregate statistic of the actions taken by all individuals is then made public. We assume that the aggregate statistic can be expressed as a function of the individuals' posterior probabilities of some event. We then show (Theorem 1) that if the aggregation function satisfies a condition we call "stochastic regularity," then whenever the statistic is common knowledge, all individuals must agree on the posterior probabilities of the event.

We next show (Theorem 2) that with finite information partitions, an iterative process of public announcement of the statistical information converges to the situation of common knowledge described above. The iterative process consists of a series of periods, where, in each period, individuals compute posterior probabilities and act on the basis of their current information. Then

the aggregate statistical information is announced, individuals recompute posterior probabilities based on their previous information <u>plus</u> the new public information, and so on.

Finally, we show (Theorem 3), that there is a close connection between a common knowledge situation and a rational expectations equilibrium. Thus, our results have implications for certain rational expectations models. In particular, in market models, if the Walrasian correspondence satisfies our regularity condition, then the rational expectations equilibrium is one of consensus, whether or not it is fully revealing. Further, the iterative process of Theorem 2 can be interpreted as a mechanism by which individuals gradually refine and learn the state-price correspondence over time.

The result of this paper generalizes results of Aumann [1978] and Geanakoplos and Polemarchakis [1982]. Aumann shows that if two individuals have the same prior and their posteriors of some event are common knowledge, then their posteriors must be equal. Geanakoplos and Polemarchakis [1982] show that if two individuals start with the same priors and then successively report their posterior probabilities of some event to each other, then eventually this leads to a situation of common knowledge where their posterior probabilities must be equal. This paper is more general than the cited papers in that we consider n individuals, and we only require an aggregate statistic of the posterior probabilities, rather than the posterior probabilities of each individual, to be common knowledge.

Several examples illustrate the problem of information aggregation that we will be addressing. These examples are developed in more detail in Section 7.

Example 1 (A Delphi Process) A panel of experts, each with his own private information, must, as a group, make a prediction on the likelihood of a certain event (i.e., the likelihood of a nuclear accident, the likelihood of a candidate winning the election, the likelihood the defendant is guilty, etc.) The "Delphi" method has been suggested as a way of pooling the information of the different experts in such a way as to allow individual experts to benefit from the information they know others to have, while at the same time protecting the confidentiality of each individual's information and also preventing some members, by force of personality, from having an undue influence on the group (see eg., Dalkey [1969]). The Delphi method works as follows: The experts are isolated from one another and each expert makes his own prediction of the likelihood based on his private information only. The average, or some other statistic of these predictions is then announced to all the experts. They then make new predictions, based on their original information plus whatever they learn from the public information. This continues until there is no further revision in any individual predictions.

Example 2 (Decentralized Risk Assessment with Cournot Oligopolists) An industry, consisting of n chemical firms is preparing to market a new chemical. Each firm has done its own private preliminary testing of the chemical's toxicity, obtaining information it does not share

with its competitors. Definitive tests on the toxicity of the chemical are being conducted by an independent firm, and will not be available for several years. In the meantime, firms may produce as much of the chemical as they wish, but if the chemical is eventually found to be toxic, the firm may expect to pay liability in proportion to its total production. In each production period, the firms have access to data on the total sales volume or market price of the chemical.

Example 3 (Parimutuel Betting) Bettors at a racetrack can bet on one of J mutually exclusive events—i.e., on which horse comes in first. Each bettor comes to the race with some private information, but before he bets, he has access to aggregate data on the information of others, in the form of the odds of each horse, which are posted on the totalizer.

Example 4 (Markets with Incomplete Information) Grossman [1981] describes an example of an economy, with incomplete information, in which there is a complete set of state contingent futures markets. Each consumer begins with some private information, and, given any price vector, chooses state contingent consumption. In addition to his private information, each individual observes the vector of equilibrium prices.

We show, in Section 7, that in all of these examples, with appropriately chosen preferences, the public information satisfies our regularity conditions. Hence, in each example, if the public information is common knowledge, there is consensus in beliefs. Further, an iterative process of public announcement will lead to this common knowledge situation. Finally, a rational expectations equilibrium to each example is also characterized by consensus.

2. Common Knowledge

Let $(\Omega, \underline{F}, \rho)$ be a probability space, and $N = \{1,2,\ldots,n\}$ be a set of individuals. For each $i \in N$, we assume there is a partition $\underline{P}_i$ of Ω. We write $\underline{P} = (\underline{P}_1,\ldots,\underline{P}_n)$ for the n tuple of partitions. For any $\omega \in \Omega$, and $i \in N$, let $P_i(\omega)$ denote the element of $\underline{P}_i$ containing ω. We assume $P_i(\omega) \in \underline{F}$ for all $i \in N, \omega \in \Omega$, and write $\underline{P}_i \subseteq \underline{F}$ for the σ-algebra generated by $\underline{P}_i$.

To define i's posterior probability of some event $A \in \underline{F}$, we fix $A \in \underline{F}$, and fix a version $\rho(A|\underline{P}_i):\Omega \to \mathbb{Z}$ of the conditional probability of A given $\underline{P}_i$. For any state $\omega \in \Omega$, and agent $i \in N$, we let $\rho(A|P_i(\omega))$ be the conditional probability of A given $\underline{P}_i$, evaluated at ω. Thus, $\rho(A|P_i(\omega))$ is the posterior probability under $\underline{P}_i$ of agent i for A. So, for all $P \in \underline{P}_i$,

$$\int_P \rho(A|P_i(\omega))d\rho(\omega) = \rho(A \cap P). \tag{2.1}$$

We define $\hat{\underline{P}} = \underline{P}_1 \wedge \underline{P}_2 \wedge \cdots \wedge \underline{P}_n$ to be the meet of $\underline{P}_1,\ldots,\underline{P}_n$, and $\tilde{\underline{P}} = \underline{P}_1 \vee \underline{P}_2 \vee \cdots \vee \underline{P}_n$ to be the join of $\underline{P}_1,\ldots,\underline{P}_n$.[1] For any $\omega \in \Omega$, we let $\hat{P}(\omega)$ denote the element of $\hat{\underline{P}}$ containing ω and $\tilde{P}(\omega)$

denote the element of $\tilde{\underline{P}}$ containing ω. Let $\hat{\underline{P}}$ and $\tilde{\underline{P}}$ be the σ-algebras generated by $\hat{\underline{P}}$ and $\tilde{\underline{P}}$ respectively.

Following Aumann [1976], we say an event $E \subseteq \Omega$ is <u>common knowledge at</u> ω <u>under</u> $\underline{P}$ iff $\hat{P}(\omega) \subseteq E$. Aumann shows that this definition of common knowledge is equivalent to the following infinite sequence of conditions: Each individual knows E, each knows all others know E, each knows that all know that all know E, etc.[1]

We extend Aumann's idea to define when the value of a statistic is common knowledge. We let Δ be an arbitrary set, and let $\varphi: \Omega \to \Delta$ be a statistic. Then we say that φ is <u>common knowledge at</u> ω <u>under</u> $\underline{P}$ iff $\hat{P}(\omega) \subseteq \{\omega' \in \Omega | \varphi(\omega') = \varphi(\omega)\}$. So φ is common knowledge at ω iff the inverse image of $\varphi(\omega)$ is common knowledge at ω under $\underline{P}$. To say that φ is common knowledge at ω is to say that once $\omega \in \Omega$ is drawn and individuals are given their private information, each individual can infer the value of the statistic on the basis of his own information, without any further refinement of his information partition. Further each individual knows everyone knows the value of $\varphi(\omega)$, etc.

Typically, if there is any independence in the information sources giving rise to the individual partitions, the elements $\hat{P}(\omega)$ of the meet will be large sets, frequently with Ω being the only member. Thus, as is argued by Geanakoplos and Polemarchakis [1982], it seems rare that an event, or some statistic φ, could be common knowledge without a process which refines individuals' initial information partitions. One way in which φ could be common knowledge is if its value were publicly announced with everyone present, when ω is drawn. Then each individual information partition (as well as the meet of all partitions) would necessarily be at least as fine as the partition generated by φ. However, common knowledge can also arise in more subtle ways, as we show in subsequent sections. Section 4 shows that an iterative process of public announcement leads to a situation of common knowledge, and Section 5 shows that a rational expectations equilibrium is a situation where the public information, φ, is common knowledge.

We will be concerned with the case when φ takes a particular form. Namely, we say φ is <u>admissible</u> if it is of the form

$$\varphi = h \circ q, \tag{2.2}$$

where $h:\mathbb{Z}^n \to \mathbb{R}$, and $q:\Omega \to \mathbb{Z}^n$ is defined by, for all $\omega \in \Omega$, $i \in N$,

$$q_i(\omega) = \rho(A|P_i(\omega)), \tag{2.3}$$

for some fixed $A \in \underline{\underline{F}}$. Thus φ is some aggregation of the individual posterior probabilities of some event. Our main theorem investigates what inferences can be made about $q(\omega)$ when φ is common knowledge. We show that if h satisfies certain conditions, then common knowledge of φ at ω implies consensus among individuals on the posterior probabilities. I.e., $q_i(\omega) = q_j(\omega)$ for all $i,j \in N$.

3. Stochastic Regularity

We now define stochastic regularity, the condition on the information aggregation function, h, which we use in our results. Let $\mathbb{Z} = \{z \in \mathbb{R} | 0 \leq z \leq 1\}$ be the unit interval and $\mathbb{Z}^n = \prod_{i=1}^{n} \mathbb{Z}$ be the unit cube in $\mathbb{R}^n$. Let $\underline{\underline{\mathbb{Z}}}^n$ be the Borel sets of $\mathbb{Z}^n$, and $\underline{M}(\mathbb{Z}^n)$ the set of probability measures on $\underline{\underline{\mathbb{Z}}}^n$. We write $\underline{\underline{\mathbb{Z}}}$ for $\underline{\underline{\mathbb{Z}}}^1$.

For any $\underline{\underline{\mathbb{Z}}}^n$ measurable $f:\mathbb{Z}^n \to \mathbb{R}$, we extend f to have domain $\underline{M}(\mathbb{Z}^n)$ by defining, for any $\lambda \in \underline{M}(\mathbb{Z}^n)$

$$f(\lambda) = E_\lambda(f(z)) = \int f d\lambda. \tag{3.1}$$

For any $\lambda \in \underline{M}(\mathbb{Z}^n)$, define the i^{th} marginal distribution of λ, denoted $\lambda_i \in \underline{M}(\mathbb{Z})$, by $\lambda_i(C) = \lambda(\{z \in \mathbb{Z}^n | z_i \in C\})$ for all $C \in \underline{\underline{\mathbb{Z}}}$.

For any $z, w \in \mathbb{Z}^n$, we write $z \geq w$ iff $z_i \geq w_i$ for all i, and we write $z > w$ iff $z \geq w$ but $z \neq w$. We say $h:\mathbb{Z}^n \to \mathbb{R}$ is <u>monotonic</u>[3] iff, for any $z,w \in \mathbb{Z}^n$, $z > w \Rightarrow h(z) > h(w)$. For any $b \in \mathbb{Z}$, let $L_b = \{z \in \mathbb{Z} | z \leq b\}$. For any $\lambda, \mu \in \underline{M}(\mathbb{Z})$, we say λ <u>stochastically dominates</u> μ, written $\lambda \geq \mu$ iff, for all $0 \leq b \leq 1$, $\lambda(L_b) \leq \mu(L_b)$. For any $\lambda, \mu \in \underline{M}(\mathbb{Z}^n)$, we write $\lambda \geq \mu$ iff $\lambda_i \geq \mu_i$ for all i, and we write $\lambda > \mu$ iff $\lambda \geq \mu$ but $\lambda \neq \mu$.

A function $f:\mathbb{Z}^n \to \mathbb{R}$ is <u>stochastically monotone</u> iff, for all $\lambda,\mu \in \underline{M}(\mathbb{Z}^n)$,

$$\lambda > \mu \Rightarrow f(\lambda) > f(\mu). \tag{3.2}$$

The function $h: \mathbb{Z}^n \to \mathbb{R}$ is <u>stochastically regular</u> iff it can be written in the form

$$h = g \circ f. \tag{3.3}$$

where $f: \mathbb{Z}^n \to \mathbb{R}$ is stochastically monotone and $g: \mathbb{R} \to \mathbb{R}$ is invertible on the range of g.

Proposition 1: If $f: \mathbb{Z}^n \to \mathbb{R}$ is additively separable into monotonic components, i.e., $f(z) = \sum_{i=1}^{n} f_i(z_i)$, where $f_i: \mathbb{Z} \to \mathbb{R}$ is monotonic, then f is stochastically monotone.

Proof: If $\lambda > \mu$, then $f(\lambda) = \int f(z)d\lambda(z) = \int(\sum_{i=1}^{n} f_i(z_i))d\lambda(z)$

$$= \sum_{i=1}^{n}\int f_i(z_i)d\lambda(z) = \sum_{i=1}^{n}\int f_i(z_i)d\lambda_i(z_i) > \sum_{i=1}^{n}\int f_i(z_i)d\mu_i(z_i)$$

$$= \sum_{i=1}^{n}\int f_i(z_i)d\mu(z) = \int(\sum_{i=1}^{n} f_i(z_i))d\mu(z) = \int f(z)d\mu(z) = f(\mu).$$

Q.E.D.

It follows if f is linear and monotonic, it is stochastically monotone.

We now give some examples of stochastically regular functions. In these examples, we let $\alpha_i \in \mathbb{R}^+$, and $G: \mathbb{R} \to \mathbb{R}$ be invertible.

Example 3.1. $h(z) = G(\sum_{i=1}^{n} \alpha_i z_i)$. Set $f(z) = \sum_{i=1}^{n} \alpha_i z_i$, and $g(t) = G(t)$.

Example 3.2 $h(z) = G(\prod_{i=1}^{n} z_i^{\alpha_i})$. Set $f(z) = \sum_{i=1}^{n} \alpha_i \log z_i$ and $g(t) = G(e^t)$.

Example 3.3 $h(z) = G(||z||)$. Set $f(z) = \sum_{i=1}^{n} z_i^2$ and $g(t) = G(t^{1/2})$.

4. A Theorem on Common Knowledge

In this section, we fix $A \in \underline{F}$, and let $q(\omega) = (q_1(\omega),\ldots,q_n(\omega))$ be defined, as in (2.3), to be the vector of posterior probabilities of the event A. We assume that $\hat{\underline{P}}$ admits a regular conditional probability, and we fix $\rho(\cdot|\hat{\underline{P}})$ to be a regular conditional probability on $\underline{F}$ given $\hat{\underline{P}}$. For any $\omega \in \Omega$, we use the notation $\rho(\cdot|\hat{P}(\omega)):\underline{F} \to \mathbb{Z}$ to denote $\rho(\cdot|\hat{\underline{P}})$ evaluated at ω.

Theorem 1. Let $\varphi = h \circ q$, where h satisfies stochastic regularity. For almost all $\omega^* \in \Omega$, if φ is common knowledge at ω^* under $\underline{P}$, then for all $i \in N$, $q_i(\omega^*) = \rho(A|\hat{P}(\omega^*))$.

Before proceeding with the proof of the Theorem, we give an informal outline of the proof for the case when each $\underline{P}_i$ is finite with $\rho(P_i(\omega) \cap A) > 0$ for all $i \in N$, $\omega \in \Omega$. In this case all conditional probabilities can be defined in the usual fashion, without recourse to regular conditional probabilities.

(a) We begin by viewing $q_i(\omega) = \rho(A|P_i(\omega))$ as a random variable with values in $\mathbb{Z}$, and hence $q(\omega)$ as a random vector with values in $\mathbb{Z}^n$. For any $\omega^* \in \Omega$, we define two conditional probability measures on the Borel sets C of $\mathbb{Z}^n$:

$$\mu(C) = \rho(\{\omega \varepsilon \Omega | q(\omega) \varepsilon C\} | \hat{P}(\omega^*)) = \rho(q^{-1}(C) | \hat{P}(\omega^*))$$

$$\lambda(C) = \rho(\{\omega \varepsilon \Omega | q(\omega) \varepsilon C\} | \hat{P}(\omega^*) \cap A) = \rho(q^{-1}(C) | \hat{P}(\omega^*) \cap A)$$

(b) The most important step, and key to the Theorem, is to observe that $\lambda > \mu$ (i.e. λ stochastically dominates μ) except for the case when $q(\omega)$ is constant on $\hat{P}(\omega^*)$, in which case $\lambda = \mu$. We discuss the intuition for this observation after step (f).

(c) An immediate consequence of the hypothesis that φ is common knowledge is that $f(\lambda) = f(\mu)$, where $h = g \circ f$.

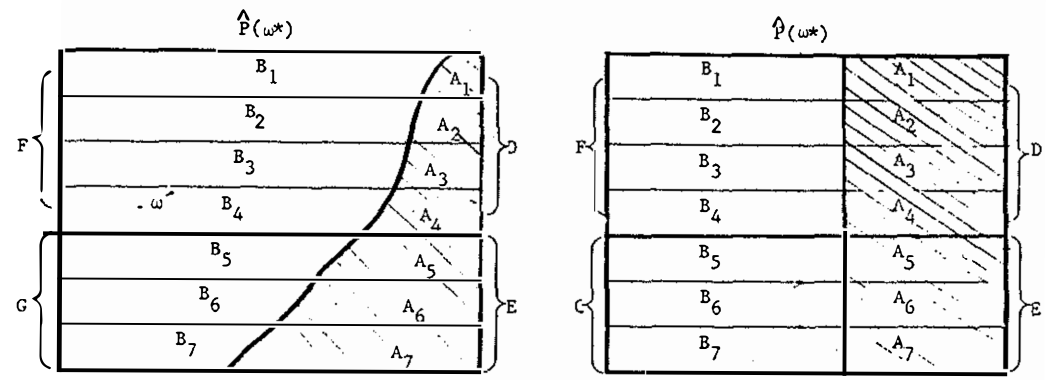
(d) By the definition of stochastic regularity of h, and stochastic monotonicity of f, we know that $\lambda > \mu \Rightarrow f(\lambda) > f(\mu)$.

(e) Putting (b), (c), and (d) together, we conclude that $\lambda = \mu$, and $q(\omega)$ is constant on $\hat{P}(\omega^*)$.

(f) Since $q(\omega)$ is constant on $\hat{P}(\omega^*)$ it follows that $q_i(\omega^*) = \rho(A | \hat{P}(\omega^*))$.

Figure 1 provides intuition for step (b). In Figure 1a, the whole rectangle represents $\hat{P}(\omega^*)$. We illustrate the case when individual i has a finite number of elements in his information partition, with j^{th} element denoted P_{ij}. So, in Figure 1a, the small rectangles represent i's information sets, P_{ij}, which, by definition of $\hat{\underline{P}}$ are subsets of $\hat{P}(\omega^*)$. We define $A_j = P_{ij} \cap A$, and $B_j = P_{ij} \cap \bar{A}$, where $\bar{A}$ is the complement of A. Thus, $P_{ij} = A_j \cup B_j$. For the particular ω' shown in Figure 1a, i's information set is $P_i(\omega') = P_{i4} = A_4 \cup B_4$. We write $a_j = \rho(A_j)$ and $b_j = \rho(B_j)$. The figure is drawn with areas proportional to these probabilities.

Figure 1a represents the case where $q_i(\omega)$ varies with



1a 1b

Figure 1

$\omega \in \hat{P}(\omega^*)$. We have drawn the figure so $q_i(\omega)$ increases as ω moves from upper to lower information sets. Note that for each j, if $\omega \in P_{ij}$, then $q_i(\omega) = a_j/(a_j + b_j)$. Writing $c_j = a_j/(a_j + b_j)$, and using the fact that $\hat{\underline{P}}$ is a coarsening of $\underline{P}_i$, we also have $\rho(A|\hat{P}(\omega^*) \cap q_i^{-1}(c_j)) = c_j$. We pick an arbitrary c_j and want to show $\lambda(L_{c_j}) < \mu(L_{c_j})$. (In Figure 1a, we pick j = 4). We write c for c_j.

Note that $\rho(A|\hat{P}(\omega^*) \cap q_i^{-1}(L_c)) = \underline{c}$ is a weighted average of conditional probabilities c_k with each $c_k \leq c$ and some strict inequalities. Hence $\underline{c} < c$. Similarly $\rho(A|\hat{P}(\omega^*) \cap q_i^{-1}(G_c)) = \bar{c} > c$, where $G_c = (c,1]$. Since $\rho(A|\hat{P}(\omega^*))$ is a weighted average of $\underline{c}$ and $\bar{c}$, it follows that

$$\rho(A|\hat{P}(\omega^*) \cap q_i^{-1}(L_c)) < \rho(A|\hat{P}(\omega^*)). \tag{4.1}$$

Expanding both sides and rearranging terms, we get

$$\rho(q_i^{-1}(L_c)|\hat{P}(\omega^*) \cap A) < \rho(q_i^{-1}(L_c)|\hat{P}(\omega^*)) \Rightarrow \lambda(L_c) < \mu(L_c), \tag{4.2}$$

which shows the stochastic dominance. Setting $d = \sum_{i=1}^{4} a_i$, $e = \sum_{i=5}^{7} a_i$, $f = \sum_{i=1}^{4} b_i$, and $g = \sum_{i=5}^{7} b_i$, to be the measures of the sets labeled D, E, F and G in Figure 1a, we can verify (4.1) by noting that $d/(d+f) < (d+e)/(d+e+f+g)$. Rearranging terms, this gets translated to $d/(d+e) < (d+f)/(d+e+f+g)$, which is precisely (4.2).

In Figure 1b, we have the case where $q_i(\omega)$ is constant for all $\omega \in \hat{P}(\omega^*)$. In this case, $\rho(A|\hat{P}(\omega^*) \cap q^{-1}(L_c)) = \rho(A|\hat{P}(\omega^*))$ so $d/(d+f) = (d+e)/(d+e+f+g)$ and $\lambda(L_c) = \mu(L_c)$.

The main difficulty in formalizing the above intuition is in defining probabilities which may be conditioned on sets of measure zero. To avoid this difficulty, our formal proof follows a slightly different route in steps (a) and (b).

<u>Proof of Theorem 1</u>.

If $\rho(A) = 0$, then for all $i \in N$, a.e. $\omega' \in \Omega$ and a.e. $\omega^* \in \Omega$, we have $q_i(\omega') = 0 = \rho(A|\hat{P}(\omega^*))$, and the result is trivially true. So assume $\rho(A) \neq 0$.

(a) We define the measure $\rho_A:\underline{\underline{F}} \to \mathbb{R}$ by, for all $P \in \underline{\underline{F}}$, $\rho_A(P) = \rho(A \cap P)/\rho(A)$. Then let $\underline{\underline{P}}^*$ be the σ-algebra generated by taking all $P \in \Omega$ such that $P = q^{-1}(C)$ for some $C \in \underline{\underline{Z}}^n$. It follows (see Breiman, Theorem 4.34, p. 79) that we can define a regular conditional probability $\rho_A(\cdot|\hat{\underline{\underline{P}}})$ on $\underline{\underline{P}}^*$ given $\hat{\underline{\underline{P}}}$. We write $\rho_A(\cdot|\hat{P}(\omega)):\underline{\underline{P}}^* \to \mathbb{Z}$ for the conditional probability on $\underline{\underline{P}}^*$ evaluated at ω. Thus, for all $P \in \underline{\underline{P}}^*$, and all $Q \in \hat{\underline{\underline{P}}}$, $\int_Q \rho_A(P|\hat{P}(\omega))d\rho_A(\omega) = \rho_A(P \cap Q)$. Equivalently, for all $P \in \underline{\underline{P}}^*$ and $Q \in \hat{\underline{\underline{P}}}$

$$\int_{Q \cap A} \rho_A(P|\hat{P}(\omega^*))d\rho(\omega) = \rho(P \cap Q \cap A). \tag{4.3}$$

For any $\omega^* \in \Omega$, define the probability measures $\lambda, \mu \in \underline{M}(\mathbb{Z}^n)$ by, for all $C \in \underline{\underline{Z}}^n$,

$$\lambda(C) = \rho_A(q^{-1}(C)|\hat{P}(\omega^*))$$

and (4.4)

$$\mu(C) = \rho(q^{-1}(C)|\hat{P}(\omega^*))$$

(b) From Lemma 2, it follows that for almost all $\omega^* \in \Omega$, all $C \in \underline{\underline{Z}}$, and all $i \in N$,

$$\rho(A|\hat{P}(\omega^*))\ \lambda_i(C) = \int_C t\ d\mu_i(t). \tag{4.5}$$

But using (4.5), we can prove that for almost all $\omega^* \in \Omega$, that $\lambda \geq \mu$, with

$$\lambda > \mu \text{ unless } q(\omega) = q(\omega') \text{ for } \rho(\cdot|\hat{P}(\omega^*)) \tag{4.6}$$
$$\text{a.e. } \omega,\ \omega' \in \hat{P}(\omega^*).$$

To see (4.6), there are two cases. If $\rho(A|\hat{P}(\omega^*)) = 0$, then $\int_C t\ d\mu_i(t) = 0$ for all $C \in \underline{\underline{Z}}$ and all $i \in N \Rightarrow \mu_i(\{0\}) = 1$ for all $i \in N$. But then if $\underline{0} = (0,\ldots,0) \in \mathbb{Z}^n$, we have $\mu(\{\underline{0}\}) = 1$ so clearly $\lambda \geq \mu$. Also (4.5) holds since $\mu(\{\underline{0}\}) = 1 \Rightarrow q(\omega) = q(\omega') = 0$ for $\rho(\cdot|\hat{P}(\omega^*))$ a.e. $\omega,\omega' \in \hat{P}(\omega^*)$. On the other hand, if $\rho(A|\hat{P}(\omega^*)) \neq 0$, then we apply Lemma 1 directly to get (4.6).

(c) Next, if $\omega^* \in \Omega$ satisfies $\hat{P}(\omega^*) \subseteq \{\omega \in \Omega | h(q(\omega)) = h(q(\omega^*))\}$, then, since $h = g\circ f$, and $g:\mathbb{R} \to \mathbb{R}$ is invertible on $f(\mathbb{Z}^n)$, we have

$$\forall\ \omega \in \hat{P}(\omega^*),\ f(q(\omega)) = f(q(\omega^*)).$$

But this implies

$$\begin{aligned} f(\lambda) &= \int f(x)\ d\lambda(x) = \int f(q(\omega))\ d\rho_A(\omega|\hat{P}(\omega^*)) \\ &= f(q(\omega^*))\int d\rho_A(\omega|\hat{P}(\omega^*)) \\ &= f(q(\omega^*))\int d\rho(\omega|\hat{P}(\omega^*)) \\ &= \int f(x)\ d\mu(x) = f(\mu). \end{aligned}$$

(d) By the definition of stochastic regularity of h, we have $\lambda > \mu \Rightarrow f(\lambda) > f(\mu)$.

(e) From (b), (c), and (d) we conclude $\lambda = \mu$ and $q(\omega)$ is constant for $\rho(\cdot|\hat{P}(\omega^*))$ a.e. $\omega \in \hat{P}(\omega^*)$.

(f) But then, using this result together with (4.5), we have for ρ a.e. $\omega^* \in \Omega$, if $\hat{P}(\omega^*) \subseteq \{\omega \in \Omega | h(q(\omega)) = h(q(\omega^*))\}$, then for all $i \in N$, there is a set W of $\rho(\cdot|\hat{P}(\omega^*))$ measure 1 s.t. for any $\omega' \in W$,

$$\begin{aligned} \rho(A|\hat{P}(\omega^*))\lambda_i(\mathbb{Z}) &= \int_{\mathbb{Z}} t\ d\mu_i(t) \\ &= \int_{\hat{P}(\omega^*)} q_i(\omega)\ d\rho(\omega|\hat{P}(\omega)^*)) \\ &= q_i(\omega') \int_{\hat{P}(\omega^*)} d\rho(\omega|\hat{P}(\omega^*)) = q_i(\omega') \end{aligned}$$

So $\rho(A|\hat{P}(\omega^*)) = q_i(\omega')$ for all $\omega' \in H$. I.e., for $\rho(\cdot|\hat{P}(\omega^*))$ a.e. $\omega \in \hat{P}(\omega^*)$, $q_i(\omega) = \rho(A|\hat{P}(\omega^*))$. The result now follows using the fact that any set of $\rho(\cdot|\hat{P}(\omega^*))$ measure zero, for ρ almost all ω^*, is also of ρ measure zero.

(Q.E.D.)

<u>Lemma 1</u>. Let $\lambda,\mu \in \underline{M}(\mathbb{Z})$, and assume there is a monotonic function $\emptyset:\mathbb{Z} \to \mathbb{Z}$ such that, for all $C \in \underline{\underline{Z}}$, $\lambda(C) = \int_C \emptyset\ d\mu$. Then $\lambda \geq \mu$.

with $\lambda > \mu$ unless $\mu(\{t\}) = 1$ for some $t \in \mathbb{Z}$.

Proof: For any $c \in \mathbb{Z}$, write $L_c = [0,c]$, and $G_c = (c,1]$. Since λ and μ are both probability measures, we must have $\mu(\mathbb{Z}) = \lambda(\mathbb{Z}) = \int_{\mathbb{Z}} \phi \, d\mu = 1$. So for $c \in \mathbb{Z}$,

$$\lambda(L_c) \leq \mu(L_c) \iff \int_{L_c} \phi \, d\mu \leq \mu(L_c)$$
$$\iff [\mu(L_c) + \mu(G_c)]\int_{L_c} \phi \, d\mu \leq \mu(L_c)[\int_{L_c} \phi \, d\mu + \int_{G_c} \phi \, d\mu]$$
$$\iff \mu(G_c)\int_{L_c} \phi \, d\mu \leq \mu(L_c)\int_{G_c} \phi \, d\mu. \tag{4.7}$$

Now, if $\mu(L_c)\mu(G_c) = 0$, then both sides of the above inequality are 0 (since $\mu(L_c) = 0 \Rightarrow \int_{L_c} \phi \, d\mu = 0$ and $\mu(G_c) = 0 \Rightarrow \int_{G_c} \phi \, d\mu = 0$). So in this case (4.7) is an equality. If $\mu(L_c)\mu(G_c) \neq 0$, however, then (4.7) is true iff

$$\frac{\int_{L_c} \phi \, d\mu}{\mu(L_c)} \leq \frac{\int_{G_c} \phi \, d\mu}{\mu(G_c)}$$
$$\iff E(\phi|L_c) \leq E(\phi|G_c). \tag{4.8}$$

Since ϕ is monotone, we have

$$E(\phi|L_c) < \phi(c) < E(\phi|G_c).$$

So (4.8) is actually a strict inequality. Thus we have shown that for all c, $\lambda(L_c) \leq \mu(L_c)$, with equality for all c iff $\mu(L_c)\mu(G_c) = 0$ for all c. But this latter holds iff $\mu(\{t\}) = 1$ for some $t \in \mathbb{Z}$. This shows $\lambda \geq \mu$ with $\lambda > \mu$ unless $\mu(\{t\}) = 1$ for some $t \in \mathbb{Z}$.

Q.E.D.

Lemma 2. For a.e. $\omega^* \in \Omega$, all $C \in \underline{\underline{Z}}$, and all $i \in N$,

$$\rho(A|\hat{P}(\omega^*))\lambda_i(C) = \int_C t \, d\mu_i(t).$$

Proof: We must show that for almost all $\omega^* \in \Omega$ and all $C \in \underline{\underline{Z}}$,

$$\rho(A|\hat{P}(\omega^*))\rho_A(q_i^{-1}(C)|\hat{P}(\omega^*)) = \int_{q^{-1}(C)} q_i(\omega) \, d\rho(\omega|\hat{P}(\omega^*)).$$

It suffices to show that for all $Q \in \hat{\underline{\underline{P}}}$,

$$\int_Q (\omega)\rho_A(q_i^{-1}(C)|\hat{P}(\omega^*))\rho(A|\hat{P}(\omega^*))\rho(d\omega^*)$$
$$= \int_Q [\int_{q^{-1}(C)} q_i(\omega)\rho(d\omega|\hat{P}(\omega^*))]\rho(d\omega^*). \tag{4.9}$$

But using the fact that $\rho(\cdot|\hat{\underline{\underline{P}}})$ is a regular conditional probability, and applying (4.3), we can write the left hand side of (4.9) as

$$\int_Q \rho_A(q_i^{-1}(C)|\hat{P}(\omega^*))[\int_A \rho(d\omega|\hat{P}(\omega^*))]\rho(d\omega^*)$$

$$= \int_Q \int_A \rho_A(q_i^{-1}(C)|\hat{P}(\omega^*))\rho(d\omega|\hat{P}(\omega^*))\rho(d\omega^*)$$

$$= \int_{Q \cap A} \rho_A(q_i^{-1}(C)|\hat{P}(\omega^*))\rho(d\omega^*)$$

$$= \rho(q_i^{-1}(C) \cap A \cap Q).$$

Since $\hat{\underline{\underline{P}}}$ is a coarsening of $\underline{P}_i$, $Q \cap q_i^{-1}(C)$ is in the σ-algebra generated by $\underline{P}_i$. So, using (2.1), the right hand side of (4.9) is

$$\int_{Q \cap q^{-1}(C)} q_i(\omega^*)\rho(d\omega^*) = \int_{Q \cap q^{-1}(C)} \rho(A|P_i(\omega^*))\rho(d\omega^*)$$

$$= \rho(q_i^{-1}(C) \cap A \cap Q).$$

Q.E.D.

We provide an example in which λ and μ are explicitly calculated and the stochastic dominance of λ over μ is illustrated.

Example 4.1. Let $\Omega = \{\omega_1,\ldots,\omega_6\}$ be finite, with $\rho(\omega_i) = \frac{1}{6}$ for all i. Let $A = \{\omega_1,\omega_2,\omega_3\}$, and let $n = 3$, with information partitions as follows (we write ijk for $\{\omega_i,\omega_j,\omega_k\}$):

$$\underline{P}_1 = \{126,345\},\ \underline{P}_2 = \{135,246\},\ \underline{P}_3 = \{234,156\}.$$

Then $\hat{P}(\omega) = \Omega$ and

$$q(\omega_1) = \tfrac{1}{3}(2,2,1)$$
$$q(\omega_2) = \tfrac{1}{3}(2,1,2)$$
$$q(\omega_3) = \tfrac{1}{3}(1,2,2)$$
$$q(\omega_4) = \tfrac{1}{3}(1,1,2)$$
$$q(\omega_5) = \tfrac{1}{3}(1,2,1)$$
$$q(\omega_6) = \tfrac{1}{3}(2,1,1).$$

By definition, $\mu(C) = \rho(q^{-1}(C)|\hat{P}(\omega)) = \rho(q^{-1}(C))$ and $\mu_1(C) = \rho(q_1^{-1}(C))$. So for any $0 \le z \le 1$,

$$q_1^{-1}(\{z\}) = \begin{cases} \{\omega_3,\omega_4,\omega_5\} & \text{if } z = \frac{1}{3} \\ \{\omega_1,\omega_2,\omega_6\} & \text{if } z = \frac{2}{3} \\ \emptyset & \text{otherwise,} \end{cases}$$

and

$$\mu_1(\{z\}) = \begin{cases} \rho(\{\omega_3,\omega_4,\omega_5\}) = \frac{1}{2} & \text{if } z = \frac{1}{3} \\ \rho(\{\omega_1,\omega_2,\omega_6\}) = \frac{1}{2} & \text{if } z = \frac{2}{3} \\ \rho(\emptyset) = 0 & \text{otherwise.} \end{cases}$$

Similarly $\lambda(C) = \rho(q^{-1}(C)|A \cap \hat{P}(\omega)) = \rho(q^{-1}(C)|A)$. So

$$\lambda_1(\{z\}) = \begin{cases} \rho(\{\omega_3,\omega_4,\omega_5\}|\{\omega_1,\omega_2,\omega_3\}) = \frac{1}{3} & \text{if } z = \frac{1}{3} \\ \rho(\{\omega_1,\omega_2,\omega_6\}|\{\omega_1,\omega_2,\omega_3\}) = \frac{2}{3} & \text{if } z = \frac{2}{3} \\ \rho(\emptyset|\{\omega_1,\omega_2,\omega_3\}) = 0 & \text{otherwise.} \end{cases}$$

Note that $\lambda_1 > \mu_1$. It is easy to check that $\lambda_j = \lambda_1$ and $\mu_j = \mu_1$ for $j = 2,3$. Hence $\lambda > \mu$. In this example there are several instances in which $q_i(\omega) \neq q_j(\omega)$, instances of deviation from consensus. Step (b) of the theorem goes further to say that <u>any</u> deviation from consensus (by any individual at any $\omega \in \hat{P}(\omega^*)$) yields $\lambda > \mu$.

Of course it is possible to define, for this example, an aggregation function $h = g \circ f$ satisfying stochastic regularity. However, $\varphi = h \circ q$ cannot be common knowledge at any ω. (If φ were common knowledge, $f(q(\omega_i)) = f(q(\omega_j))$ for all $\omega_i\ \omega_j \in \hat{P}(\omega) = \Omega$, which implies $f(\lambda) = f(\mu)$. But since $\lambda > \mu$, we have by stochastic regularity $f(\lambda) > f(\mu)$, which is a contradiction.)

We show by a counter example that it is not possible to weaken the condition of stochastic regularity to one of monotonicity and still obtain the result of Theorem 1.

Example 4.2. The example is the same as Example 3.1 except we add six more individuals with information partitions:

$\underline{P}_4 = \{125,\ 346\}$, $\underline{P}_5 = \{134,\ 256\}$, $\underline{P}_6 = \{236,\ 145\}$

$\underline{P}_7 = \{124,\ 356\}$, $\underline{P}_8 = \{136,\ 245\}$, $\underline{P}_9 = \{235,\ 146\}$.

Then

$$q(\omega_1) = \tfrac{1}{3}(2,2,1,2,2,1,2,2,1)$$
$$q(\omega_2) = \tfrac{1}{3}(2,1,2,2,1,2,2,1,2)$$
$$q(\omega_3) = \tfrac{1}{3}(1,2,2,1,2,2,1,2,2)$$
$$q(\omega_4) = \tfrac{1}{3}(1,1,2,1,2,1,2,1,1)$$
$$q(\omega_5) = \tfrac{1}{3}(1,2,1,2,1,1,1,1,2)$$
$$q(\omega_6) = \tfrac{1}{3}(2,1,1,1,1,2,1,2,1).$$

It is easy to check that for no i,j do we have $q(\omega_i) > q(\omega_j)$. Thus, it is possible to construct a (continuous) monotonic function, $f\colon \mathbb{Z}^9 \to \mathbb{Z}$ such that $f(q(\omega_i)) = 0$ for all i. As before $\hat{P}(\omega_i) = \Omega$ for any i. So, setting h = f, h is common knowledge at any ω_i. But for $\omega^* = \omega_1$, we have $\rho(A|\hat{P}(\omega^*)) = \frac{1}{2} \neq \frac{2}{3} = q_1(\omega^*)$, so the result of the theorem is not true for arbitrary monotonic statistics.

We will need a slightly more general version of Theorem 1 to prove Theorem 2 below. Let $\underline{Q}$ be a coarsening of $\hat{\underline{P}}$ and $\underline{\underline{Q}}$ the σ-algebra generated by $\underline{Q}$. For the following corollary we assume $\underline{\underline{Q}}$ admits a regular conditional probability, and fix $\rho(\cdot|\underline{\underline{Q}})$ to be a regular conditional probability on $\underline{\underline{F}}$ given $\underline{\underline{Q}}$.

Corollary 1. Let $\varphi = h \circ q$, where h satisfies stochastic regularity. For any coarsening, $\underline{Q}$ of $\hat{\underline{P}}$, for almost all ω^*, ε Ω, and $\rho(\cdot|Q(\omega^*))$ a.e. $\omega' \in Q(\omega^*)$, if $Q(\omega^*) \subseteq \{\omega \in \Omega | \varphi(\omega) = \varphi(\omega^*)\}$, then $q_i(\omega') = \rho(A|Q(\omega^*))$ for all $i \in N$.

Proof: The proof is identical to that for Theorem 1, but with $Q(\omega^*)$ substituted for $\hat{P}(\omega^*)$. (Note that Lemma 1 is also true for any coarsening $\underline{Q}$ of $\hat{\underline{P}}$).

Q.E.D.

5. An Iterative Process

We begin this section with a somewhat informal discussion of an iterative process of public announcement which achieves common knowledge. We subsequently formalize the argument. This process is a generalization of the process suggested by Geanakoplos and Polemarchakis [1982], and is similar to the dynamic learning models given by Jordan [1982] and Kobayashi [1977].

Let the initial structure of information be as follows:

$(\Omega,\underline{F},\rho)$	a probability space	(5.1a)
$\underline{P}^1 = (\underline{P}^1_1,\ldots,\underline{P}^1_n)$	a collection of initial information partitions.	(5.1b)
$h\colon \mathbb{Z}^n \Rightarrow \mathbb{R}$	an aggregation function	(5.1c)

We also assume throughout this section that each individual partition $\underline{P}_i^1$ is finite and that $\rho(\tilde{P}^1(\omega)) > 0$ for each element $\tilde{P}^1(\omega)$ of the join, $\underline{\tilde{P}}^1 = \underline{P}_1^1 \vee \underline{P}_2^1 \vee \dots \vee \underline{P}_n^1$.

All individuals are informed of the structure (5.1a)-(5.1c) publicly and simultaneously so the structure of the model is itself common knowledge.[4] Nature then draws the true state, ω^*, and each i is informed of $P_i^1(\omega^*)$. Knowledge of $P_i^1(\omega^*)$ is private. Individual i calculates $q_i^1(\omega^*) = \rho(A|P_i^1(\omega^*))$ and takes some action on the basis of this posterior probability. As a consequence of all the actions an aggregation of the posterior probabilities $\varphi^1(\omega^*) = h(q^1(\omega^*))$ becomes publicly observable. Individuals use this publicly available information to refine their information sets, take new actions, and so on.

As the agents refine their information partitions in response to the public signal, it is useful to keep track of what an initially uninformed observer would infer from the process. An initially uninformed observer knows the initial structure and has access to all publicly available information, but has no initial private information. Upon observing $\varphi^1(\omega^*)$, the observer considers the possible $\omega \in \Omega$ which could have led to the observed value. The observer not only knows that the true ω must be in the inverse image of $\varphi(\omega^*)$, he also has enough information, from the general structure, to infer the inverse image set itself. We write the inverse image as

$$H^2(\omega^*) = \{\omega \in \Omega \mid h(q^1(\omega)) = h(q^1(\omega^*))\}$$

Each i knows more than the outside observer of course. Each i augments the public information with his private information and infers the true state is in $P_i^2(\omega^*) = H^2(\omega^*) \cap P_i^1(\omega^*)$. So $P_i^2(\omega^*)$ is i's second period information set, and we let $\underline{P}_i^2$ represent the partition generated by the $P_i^2(\omega)$. It is clear that $H^2(\omega^*)$ is common knowledge at ω^* under $\underline{P}^2 = (\underline{P}_1^2,\dots,\underline{P}_n^2)$. Everyone, including an outside observer, can construct $H^2(\omega^*)$. Further since $\varphi^1(\omega^*)$ is publicly announced, everyone knows the true state is in $H^2(\omega^*)$, everyone knows that everyone knows that, etc.) We can think of $H^2(\omega^*)$ as the observer's second period information set, when the true state is ω^*, and we can think of it as the second period common knowledge information set.

But the assumption that the initial information structure is common knowledge, an assumption traditionally made in the specification of Bayesian games, is a very powerful one (we did not use this assumption in Theorem 1). By means of this assumption, even before the observer sees $\varphi^1(\omega^*)$ he can infer each i's second period information partition. Since he knows $\underline{P}_i$, he knows the function $q^1(\omega)$ and hence the mapping $\underline{H}^2(\cdot)$. Thus even before observing $\varphi^1(\omega^*)$ he can construct each i's second period partition $\underline{P}_i^2 = \underline{H}^2 \vee \underline{P}_i^1$. In other words all the second period partition structures are common knowledge and can be inferred by everyone including the observer as common knowledge before nature draws ω^* and i receives his private information. This means that the inferred second period refined partitions $\underline{P}_i^2$ will be the same no matter which ω is the true state. But of course the specific private information sets $P_i^1(\omega^*)$ and $P_i^2(\omega^*)$ vary with the true state and are unknown to the observer.

After each i learns that the true state is in $P_i^2(\omega^*)$, he can compute the posterior probability of A, $q_i^2(\omega^*) = \rho(A|P_i^2(\omega^*))$. He then takes his second action, and after this, $\varphi^2(\omega^*) = h(q^2(\omega^*))$ is revealed. The observer can use his knowledge of the $\underline{P}_2^i$ to compute $q_i^2(\omega) = \rho(A|P_i^2(\omega))$ for any $\omega \in \Omega$, and hence he can compute the inverse image of $\varphi^2(\omega^*)$. Thus, the observer makes the inference that the true state is in

$$H^3(\omega^*) = \{\omega \in H^2(\omega^*) | h(q^2(\omega)) = h(q^2(\omega^*))\}.$$

In a similar fashion, define $\underline{P}_i^3 = \underline{H}^3 \vee \underline{P}_i^2$. Again $H^3(\omega^*)$ is common knowledge, at ω^* under $\underline{P}^3 = (\underline{P}_1^3,\ldots,\underline{P}_n^3)$, and so on.

The process is formally defined as follows: For all $\omega \in \Omega$, and all $i \in N$, define $H^1(\omega) = \Omega$, and let $\underline{P}^1$ be as in (5.1b). Then, inductively on t, for each $i \in N$, define for any $\omega \in \Omega$,

$$q_i^t(\omega) = \rho(A|P_i^t(\omega)), \qquad (5.2a)$$

$$q^t(\omega) = (q_1^t(\omega),\ldots,q_n^t(\omega)),$$

and, for any $\omega \in \Omega$, define

$$H^{t+1}(\omega) = \{\omega' \in H^t(\omega) | h(q^t(\omega')) = h(q^t(\omega))\}, \qquad (5.2b)$$

$$\underline{H}^{t+1} = \{H^{t+1}(\omega) | \omega \in \Omega\}, \qquad (5.2c)$$

$$\underline{P}_i^{t+1} = \underline{H}^{t+1} \vee \underline{P}_i^t = \underline{H}^{t+1} \vee \underline{P}_i^1. \qquad (5.2d)$$

By construction, for each ω, $H^1(\omega) \supseteq H^2(\omega) \supseteq \ldots \supseteq H^t(\omega)$. So $\underline{H}^{t+1}$ is a refinement of $\underline{H}^t$. As long as $H^{t+1}(\omega)$ is a proper subset of $H^t(\omega)$, the outside observer is learning, and the fund of common knowledge is improving. (It is possible to construct examples, similar to those of Geanakoplos and Polemarchakis, for which the public signal is constant for many periods but $H^t(\omega)$ is getting smaller and learning continues.) If, for some T, $\underline{H}^{T+1} = \underline{H}^T$, then for all $t > T$, $\underline{P}_i^{t+1} = \underline{P}_i^t$. This means that there is no further refinement and the process stops.

We say the process is in a <u>common knowledge equilibrium</u> if the current statistic is common knowledge relative to the <u>current</u> information sets, i.e., if φ^t is common knowledge under $\underline{P}^t = (\underline{P}_1^t,\ldots,\underline{P}_n^t)$. The common sense of this definition is that the process has reached such a point of information refinement that the statistic can be inferred before it is observed, and thus its observation leads to no further refinement in the information sets. Our next theorem shows that under the assumptions made, we will eventually reach a common knowledge equilibrium. If the statistic, $\varphi = h \circ q$ is such that h satisfies stochastic regularity, the common knowledge equilibrium is characterized by consensus. For this theorem, we write $\hat{\underline{P}}^t = \underline{P}_1^t \wedge \ldots \wedge \underline{P}_n^t$ as the meet of the period t partitions, and $\hat{P}^t(\omega)$ for the element of the meet containing ω.

<u>Theorem 2</u>. Assume an initial information structure as in (5.1), the iterative process of (5.2), and let $\varphi^t = h \circ q^t$. Then for all $\omega \in \Omega$, there is a T such that φ^T is common knowledge at ω under $(\underline{P}_1^T,\ldots,\underline{P}_n^T)$. Further, if h satisfies stochastic regularity, then, for any such T, and for all $\omega \in \Omega$, $q_i^T(\omega) = \rho(A|H^T(\omega)) = \rho(A|\hat{P}^T(\omega))$.

<u>Proof</u>: We know that for all t, $\underline{H}^{t+1}$ is a refinement of $\underline{H}^t$. Further, each $\underline{H}^t$ is a coarsening of $\underline{\tilde{P}}$. By finiteness of the $\underline{P}_i$, it follows that $\underline{\tilde{P}}$, and hence the $\underline{H}^t$ are finite. So there is a T for which $\underline{H}^{T+1} = \underline{H}^T$. For this T and for any $\omega \in \Omega$ we will show $\{\omega' \in \Omega | h(q^T(\omega')) = h(q^T(\omega))\} \supseteq \hat{P}^T(\omega)$. Note $\{\omega' \in \Omega | h(q^T(\omega')) = h(q^T(\omega))\} \supseteq \{\omega' \in H^T(\omega) | (h(q^T(\omega')) = h(q^T(\omega))\}$ $= H^{T+1}(\omega) = H^T(\omega)$. So all we have to show is that $H^T(\omega) \supseteq \hat{P}^T(\omega)$. But for each $\omega \in \Omega$ and each $i \in N$, $P_i^T(\omega) = P_i^{T-1}(\omega) \cap H^T(\omega) \subseteq H^T(\omega)$ so $\underline{H}^T$ is a common coarsening of all the $\underline{P}_i^T$ and thus $H^T(\omega) \supseteq \hat{P}^T(\omega)$. The second assertion of the Theorem now follows directly from Corollary 1, setting $\underline{Q} = \underline{H}^T$, and letting the posterior probabilities $q(\omega)$ in the corollary be computed with respect to the partitions $(\underline{P}_1^T,\ldots,\underline{P}_n^T)$. The last part of the equality follows setting $\underline{Q} = \hat{\underline{P}}^T$.

Q.E.D.

We now give an example illustrating Theorem 2.

<u>Example 5.1</u> Let $\Omega = \{1,2,3,4,5\}$, $n = 3$,

$$\underline{P}_1 = \{\{1,2,4\},\ \{3,5\}\}$$
$$\underline{P}_2 = \{\{1,3,5\},\ \{2,4\}\}$$
$$\underline{P}_3 = \{\{1,5\},\{2,3,4\}\}.$$

Also, let $A = \{1,2,3\}$, $\omega = 1$, and $h(z) = (z_1 + z_2 + z_3)/3$. Then it is easily checked that

$$q^1(\omega) = (\tfrac{2}{3},\tfrac{2}{3},\tfrac{1}{2}),\ h(q^1(\omega)) = \tfrac{11}{18},\ H^1(\omega) = \{1,2,3,4\},$$
$$q^2(\omega) = (\tfrac{2}{3},1,1),\ h(q^2(\omega)) = \tfrac{8}{9},\ H^2(\omega) = \{1,3\},$$
$$q^3(\omega) = (1,1,1),\ h(q^3(\omega)) = 1,\ H^3(\omega) = \{1,3\}.$$

So the iterative process converges with t = 3, at which point all individuals have the same posterior probabilities for A. Note that in this example, the posterior probabilities are the same as they would be under pooled information, even though the pooled information, $\{\omega\} = \bigcap_i P_i(\omega) = \tilde{P}(\omega)$, is not common knowledge.

Although all posterior probabilities must agree when the statistic satisfies stochastic regularity and ϕ becomes common knowledge, these posteriors in general need not be the same as those from pooled information. In other words we need not have $\rho(A|\hat{P}(\omega^*)) = P(A|\tilde{P}(\omega^*))$. The same example as that used by Geanakoplos and Polemarchakis [1982, p. 198] serves to illustrate this point. Also, it is possible to construct examples for n individuals where the iterative process takes any finite number of steps to converge, and it is possible to construct examples where there is no evident revision in individual posteriors until the last step. However, we conjecture that if the event A is chosen randomly from a non atomic state space, as in Proposition 4 of Geanakoplos and Polemarchakis, that the iterative process will converge, with probability one, in less than or equal to n iterations, to the pooled information situation (in Kobayashi's [1977] model convergence takes n steps).

6. Rational Expectations and Common Knowledge

The idea underlying rational expectations equilibria is that each individual has a model, or belief, of the form of the public information. In equilibrium, these beliefs are "rational" in the sense that when individuals act on them, the beliefs end up not being contradicted. We now show the connection between the common knowledge approach and the idea of rational expectations.

In the previous development, we have suppressed the dependence of the public information on the information partition. The statistics we have considered are of the form $\varphi:\Omega \rightarrow \Delta$, where $\Delta = \mathbb{R}$ and $\varphi(\omega) = h\circ q(\omega) = h(\rho(A|P_1(\omega)),\ldots,\rho(A|P_n(\omega)))$. For arbitrary partitions $\underline{R} = (\underline{R}_1,\ldots,\underline{R}_n)$, we can express the dependence of φ on $\underline{R}$ by writing

$$\varphi(\omega,\underline{R}) = h\circ q(\omega,\underline{R}) = h(\rho(A|R_1(\omega)),\ldots,\rho(A|R_n(\omega))) \tag{6.1}$$

where $q_i(\omega,R) = \rho(A|R_i(\omega))$. In general $\varphi(\omega,\underline{R})$ could be an arbitrary function which need not take the particular from above. For example, if $\varphi(\omega,\underline{R})$ represents the Walrasian equilibrium price correspondence, in general it would depend on the whole posterior probability

distributions, $\rho(\cdot|R_i(\omega))$, rather than just on the posterior probabilities of a given event.

In a rational expectations equilibrium, it is assumed each individual has a model, $\phi:\Omega \rightarrow \Delta$ of $\varphi(\omega,\underline{R})$, and uses $\phi(\omega)$ together with his private information to condition his beliefs and actions. Thus, for $\omega \in \Omega$, let $H_\phi(\omega) = \{\omega' \in \Omega|\phi(\omega') = \phi(\omega)\}$ and $\underline{H}_\phi = \{H_\phi(\omega)|\omega \in \Omega\}$. We write $\underline{P}_{\phi,i} = \underline{P}_i \vee \underline{H}_\phi$ for the join of $\underline{P}_i$ and $\underline{H}_\phi$, with elements of the form $P_{\phi,i}(\omega) = P_i(\omega) \cap H_\phi(\omega)$, and write $\underline{P}_\phi = (\underline{P}_1 \vee \underline{H}_\phi,\ldots,\underline{P}_n \vee \underline{H}_\phi)$.

We can now define a Rational Expectations Equilibrium (REE) under $\underline{P}$ to be a $\tilde{\underline{P}}$ measurable function $\phi:\Omega \rightarrow \Delta$ which satisfies

$$\varphi(\omega,\underline{P}_\phi) = \phi(\omega) \tag{6.2}$$

for all $\omega \in \Omega$. So in a rational expectations equilibrium, if individuals act on their beliefs, the beliefs are confirmed. If ϕ is a REE under $\underline{P}$ and $\underline{P}_\phi = (\tilde{\underline{P}},\tilde{\underline{P}},\ldots,\tilde{\underline{P}})$, then ϕ is called a Fully Revealing REE. So in a fully revealing REE, each individual ends up obtaining all the information available to everyone.

The following theorem shows the close connection between the notion of common knowledge and a REE. In this proposition, when we say ϕ is common knowledge under $\underline{R} = (\underline{R}_1,\ldots,\underline{R}_n)$, we mean that it is common knowledge at ω under $\underline{R}$ for all $\omega \in \Omega$.

Theorem 3 Let $\phi: \Omega \rightarrow \Delta$. Then

(a) ϕ is a REE under $\underline{P} \Rightarrow \phi$ is common knowledge under $\underline{P}_\phi$.

(b) If ϕ is defined by $\phi(\omega) = \varphi(\omega,\underline{P})$, then ϕ is common knowledge under $\underline{P}$ => ϕ is a REE under $\underline{P}$.

Proof:

(a) Let ϕ be a REE under $\underline{P}$. So $\phi(\omega) = \varphi(\omega,\underline{P}_\phi)$ for all $\omega \in \Omega$. We must show that for any $\omega \in \Omega$, $\hat{\underline{P}}_\phi(\omega) \subseteq \{\omega' \in \Omega | \phi(\omega') = \phi(\omega)\} = H_\phi(\omega)$. But this follows from the fact that $\underline{P}_{\phi,i} = \underline{P}_i \vee \underline{H}_\phi$, so that H_ϕ is a common coarsening of the $\underline{P}_{\phi,i}$.

(b) Let $\phi(\omega) = \varphi(\omega,\underline{P})$ be common knowledge under $\underline{P}$. Then for all $\omega \in \Omega$, $\hat{\underline{P}}(\omega) \subseteq \{\omega' \in \Omega | \phi(\omega') = \phi(\omega)\} = H_\phi(\omega)$. So $P_i(\omega) \cap H_\phi(\omega) = P_i(\omega)$ for all $\omega \in \Omega$. Hence $\underline{P}_\phi = \underline{P}$, so $\phi(\omega) = \varphi(\omega,\underline{P}) = \varphi(\omega,\underline{P}_\phi)$, so ϕ is a REE under P.

Q.E.D.

For the following corollaries, we let $\underline{A} = \{A_j\}_{j=1}^J$ be a fixed collection of sets with $A_j \subseteq \Omega$ for each j. Assume that for each n-tuple of partitions, $\underline{R} = (\underline{R}_1,\ldots,\underline{R}_n)$, that $\varphi(\cdot,\underline{R}):\Omega \to \Delta$ has range $\Delta = \mathbb{R}^J$, and write $\varphi_j(\omega,\underline{R})$ for the j^{th} component of $\varphi(\omega,\underline{R})$. Also, write $q^j(\omega,\underline{R})$ for the vector with i^{th} component $q_i^j(\omega,\underline{R}) = \rho(A_j|R_i(\omega))$. The following corollary gives conditions on $\varphi(\omega,\underline{R})$ guaranteeing that every REE will achieve consensus.

Corollary 2 Let $\varphi(\cdot,\underline{R}):\Omega \to \mathbb{R}^J$ satisfy $\varphi_j(\omega,\underline{R}) = h \circ q^j(\omega,\underline{R})$ for all $\underline{R}$, all $\omega \in \Omega$, and all $1 \le j \le J$. Assume h is stochastically regular. If $\phi:\Omega \to \mathbb{R}^J$ is a REE under $\underline{P}$, then for a.e. $\omega \in \Omega$, $q_i^j(\omega,\underline{P}_\phi) = q_k^j(\omega,\underline{P}_\phi) = \rho(A_j|\hat{\underline{P}}_\phi(\omega))$ for all i,k,j.

Proof: Since ϕ is a REE under $\underline{P}$, we have $\phi(\omega) = \varphi(\omega,\underline{P}_\phi)$ for all $\omega \in \Omega$. By Theorem 3, ϕ is common knowledge under $\underline{P}_\phi$. Hence $\phi_j(\omega) = \varphi_j(\omega,\underline{P}_\phi)$ is common knowledge at ω under $\underline{P}_\phi$ for all $1 \le j \le J$. The result is now a direct application of Theorem 1.

Q.E.D.

The final corollary gives a condition on $\underline{A} = \{A_j\}_{j=1}^J$ guaranteeing that every REE is fully revealing. The condition is that the events A_j lead to a partition which is at least as fine as the pooled information. We conjecture that there are much more general conditions than those given in this corollary which yield the same result. In fact we conjecture that such a result is generically true. However the condition here is of some interest as it corresponds to the information structures which have been used in experimental work on rational expectations equilibria (see, eg., Plott and Sunder [1983]).

Corollary 3 Assume $\underline{A}$ is a partition, that $\underline{A}$ is a refinement of $\tilde{\underline{P}}$, and that $\rho(A_j) > 0$ for all j. Assume $\varphi(\cdot,\underline{R})$ satisfies the conditions of Corollary 2. Then if $\phi:\Omega \to \mathbb{R}^J$ is a REE under $\underline{P}$, it is a fully revealing REE.

Proof: As above, we have ϕ is common knowledge under $\underline{P}_\phi$. For convenience, write $\underline{P}_{\phi,i} = \underline{R}_i$ for all i. Now pick $\omega^* \in \Omega$. We prove that $R_i(\omega^*) = R_k(\omega^*)$ for all i, k $\in$ N. Assume not, and assume, w.l.o.g., that $\omega' \in R_i(\omega^*) - R_k(\omega^*)$. Then $R_k(\omega') \cap R_k(\omega^*) = \emptyset$. But since $\underline{A}$ is a partition and it refines $\tilde{\underline{P}}$, we can pick $A_j \in \underline{A}$ with $A_j \subsetneq \tilde{R}(\omega^*)$, which implies $A_j \subsetneq R_k(\omega^*)$. So $A_j \cap R_k(\omega') = \emptyset$, and hence $\rho(A_j|R_k(\omega')) = 0$. Now $\omega' \in \hat{P}(\omega^*)$. So by Theorem 1, we have

$$\rho(A_j|R_k(\omega^*)) = \rho(A_j|R_i(\omega')) = \rho(A_j|R_i(\omega^*)) = \rho(A_j|R_k(\omega')) = 0$$

But $\rho(A_j|R_k(\omega^*)) = \rho(A_j)/\rho(R_k(\omega^*)) \neq 0$, a contradiction. But then $R_i(\omega^*) = R_k(\omega^*)$ for all i, k $\in N$. I.e. $\underline{P}_{\emptyset,i} = \underline{P}_{\emptyset,k} = \tilde{\underline{P}}_\emptyset$ for all i,k. But each $\underline{P}_{\emptyset,i}$ is a refinement of $\underline{P}_i$, and a coarsening of $\tilde{\underline{P}}$ (since $\emptyset$ is $\tilde{\underline{P}}$ measurable). Hence, $\tilde{\underline{P}}_\emptyset = \tilde{\underline{P}}$. So $\underline{P}_\emptyset = (\tilde{\underline{P}},\ldots,\tilde{\underline{P}})$.

Q.E.D.

The above results, together with our theorems of the previous sections can be used to characterize the properties of rational expectations equilibria in certain situations. The examples of the next section illustrate applications.

7. Some Examples

We now return to Examples 1-4 sketched in the introduction, to show how public information of the form assumed in our theorems arises in various settings. For this section, we assume $\Omega = S \times Y^n$, where Y is an arbitrary set from which individual private signals are drawn, and $S = \{s_1, s_2, \ldots, s_J\}$ is a finite set of "utility relevant" states of the world. So elements of Ω are written $\omega = (s,y)$, where $s \in S$ and $y = (y_1,\ldots,y_n) \in Y^n$. The individual information partitions

are then defined by

$$\underline{P}_i = \{P_i(\omega)|\omega \in \Omega\}$$

where, for $\omega = (s,y) \in \Omega$,

$$P_i(\omega) = \{\omega'=(s',y') \in \Omega | y_i' = y_i\}.$$

In addition, let X be a set from which individual i selects a decision. Given $\omega \in \Omega$, we let $x(\omega) = (x_1(\omega),\ldots,x_n(\omega)) \in X^n$ be the decisions selected by each $i \in N$ (to be described in further detail in the examples). Then in each of the following examples, we assume individuals use their private information to make decisions $x(\omega)$. A function $\varphi(\omega) = \eta \circ x(\omega)$ is then made public, where $\eta: X^n \to \mathbb{R}$ is some aggregation of individual decisions. In each case, we show that φ is of the form required by our theorems.

Example 1 A Delphi Process

In this example, each of n experts is asked to report his posterior probability of an event, and then an aggregate function of the reports is made public. We can write $S = \{0,1\}$, with $s = 1$ being the event of interest. Write $A = \{\omega \in \Omega | s = 1\}$, and $q_i(\omega) = \rho(A|P_i(\omega))$. Let $X = \mathbb{Z}$ be the set of possible reports by i. Then, assuming each expert reports truthfully, we have $x_i(\omega) = q_i(\omega)$, so $x(\omega) = q(\omega)$. We assume the public information is of the form $\varphi(\omega) = \eta \circ x(\omega)$. Clearly, if η is stochastically regular, then setting $h = \eta$, φ can be written in the form $\varphi = h \circ q$, which is of the form required by our theorems.

Note that we can <u>induce</u> truthful reporting in this example if the experts are paid for their reports using a proper scoring rule (see, eg., De Finetti [1970] or Savage [1971] for definitions of proper scoring rules). For example, if the quadratic loss proper scoring rule is used, then expert i will have a utility function $u_i: X \times S \to \mathbb{R}$ of the form

$$u_i(x_i, s) = 1 - (x - s)^2.$$

Then, given $\omega \in \Omega$, each $i \in N$ will choose $x_i \in X$ to maximize $E[u_i(x_i, s) | P_i(\omega)] = (1 - x_i^2)(1 - q_i(\omega)) + (2x_i - x_i^2) q_i(\omega)$. Since u_i is a proper scoring rule, the maximization yields $x_i(\omega) = q_i(\omega)$.

<u>Example 2 Decentralized Risk Assessment with Cournot Oligopolists</u>

In this example, firms pick levels of production of a chemical based on information obtained from private toxilogical tests, and then observe data on the market price of the chemical.

Here again, we can write $S = \{0,1\}$, with $s = 1$ indicating the chemical is toxic, $A = \{\omega \in \Omega | s = 1\}$ being the event that the chemical is toxic, and $q_i(\omega) = \rho(A | P_i(\omega))$. Let $X = \mathbb{R}^+$ be the set of possible production levels of the firm. Let $D^{-1}: \mathbb{R}^+ \to \mathbb{R}^+$ be the inverse demand function, and $C_i: \mathbb{R}^+ \to \mathbb{R}$ be firm i's cost function. Given a choice $x = (x_1, \ldots, x_n) \in X^n$ of production levels of all n firms, we can write the profit function of firm i, $\pi_i: X^n \times S \to \mathbb{R}$ by

$$\pi_i(x,s) = \begin{cases} x_i D^{-1}(\bar{x}) - C_i(x_i) & \text{if } s = 0 \\ x_i D^{-1}(\bar{x}) - C_i(x_i) - \gamma x_i & \text{if } s = 1 \end{cases}$$

where $\bar{x} = \sum_{i=1}^{n} x_i$, and $\gamma \in \mathbb{Z}$ is the proportional liability. Hence the expected profit is of the form

$$E(\pi_i(x,s) | P_i(\omega)) = x_i D^{-1}(\bar{x}) - C_i(x_i) - \gamma q_i(\omega) x_i.$$

We assume that the inverse demand is linear, of the form $D^{-1}(x) = a - bx$, with $a > 0$, $b > 0$, and that individual cost functions are quadratic, of the form $C_i(x_i) = c_i x_i^2/2 + d_i x_i + e_i$, with $c_i, d_i, e_i > 0$. It follows that if firms use only their private information, and do not try and learn from contemporaneous prices or total production of others, the unique Cournot Nash equilibrium occurs when

$$x(\omega) = A^{-1}y - \gamma A^{-1} q(\omega)$$

where A is an $n \times n$ matrix with elements

$$A_{ij} = \begin{cases} b & \text{if } i \neq j \\ 2b + c_i & \text{if } i = j, \end{cases}$$

y is an $n \times 1$ vector with elements $y_i = a - d_i$, and $x(\omega)$ and $q(\omega)$ are $n \times 1$ column vectors with elements $x_i(\omega)$ and $q_i(\omega)$, respectively. We can solve for the resultant price, $\Phi(\omega)$ by writing

$$\Phi(\omega) = \eta \circ x(\omega) = D^{-1}(v' x(\omega))$$

where $v' = (1,1,\ldots,1)$ is the $n \times 1$ vector of ones. Substituting the value of $x(\omega)$, this can be rewritten in the form

$$\Phi(\omega) = [a - bv'A^{-1}y] + b\gamma v'A^{-1}q(\omega).$$

It can be shown that the elements of the vector $v'A^{-1}$ are all positive,[5] so this is a monotonic, linear function of the $q_i(\omega)$, and hence is of the form $\varphi(\omega) = h \circ q(\omega)$ required for our theorems. Note that if demand is completely elastic (with $b = 0$), then $\bar{x}(\omega) = v'x(\omega)$ is still a statistic of the form required for our theorems.

Example 3 Parimutuel Betting

In this example, n individuals have private information, and bet on the outcome of a horse race. They then observe public information on the bets of others in the form of the odds on the totalizer. Here, the utility relevant states are the events that each horse wins. Let $s = s_j$ indicate that horse j wins, $A_j = \{\omega \in \Omega | s = s_j\}$ be the event that j wins, and write $q_i^j(\omega) = \rho(A_j | P_i(\omega))$ for i's posterior probability of horse j winning. We let $X \subseteq \mathbb{R}^{J+1}$ be the set of possible state contingent money holdings that an individual can choose from. Elements of X are of the form $x_i = (x_{i0}, x_{i1}, \ldots, x_{iJ})$, and represent the situation where i chooses a bet which leaves him x_{i0} for current consumption, and which will earn winnings of x_{ij} if state j occurs. Individual utility functions $u_i : X \times S \rightarrow \mathbb{R}$ are assumed to be of the form

$$u_i(x_i, s_j) = x_{i0} + \alpha_i \ln x_{ij},$$

where $\alpha_i > 0$ for each i. Further each $i \in N$ has initial endowment of $e_i \in \mathbb{R}$.

Now if all individuals just use their private information, and do not try and learn from the posted odds, then we can compute the equilibrium odds $\varphi(\omega) = (\varphi_1(\omega), \ldots, \varphi_J(\omega))$ for each horse as follows:

If $r = (r_1, \ldots, r_J)$ are the posted odds for horses $1, \ldots, J$, espectively, then the implicit prices for state j winnings are $j = \frac{1}{r_j + 1}$, so individual i's optimization problem is the following:

$$\max_{x_i \in X} E[u_i(x_i, s) | P_i(\omega)]$$

$$\text{s.t. } x_{i0} + \sum p_j x_{ij} \leq e_i$$

r, equivalently

$$\max_{x_i \in X} [x_{i0} + \sum_{j=1}^{J} \alpha_i q_i^j(\omega) \ln x_{ij}]$$

$$\text{s.t. } x_{i0} + \sum p_j x_{ij} \leq e_i .$$

etting up the Lagrangean, and solving, this yields

$$x_{ij}(\omega) = \frac{\alpha_i}{p_j} q_i^j(\omega)$$

Iow the total take on each horse is

$$t_j = \sum_{i=1}^{n} p_j x_{ij}(\omega),$$

and, (ignoring the track take) the odds of the totalizer are set so that $r_j = (\sum_{k \neq j} t_k)/t_j$. Equivalently,

$$p_j = \frac{t_j}{\sum_{k=1}^{J} t_k} = \frac{\sum_{i=1}^{n} p_j x_{ij}(\omega)}{\sum_{k=1}^{J} \sum_{i=1}^{n} p_k x_{ik}(\omega)} = \frac{\sum_{i=1}^{n} \alpha_i q_i^j(\omega)}{\sum_{i=1}^{n} \alpha_i \sum_{k=1}^{J} q_i^k(\omega)}$$

$$= \sum_{i=1} \left[\frac{\alpha_i}{\sum_{k=1}^{n} \alpha_k} \right] q_i^j(\omega) .$$

Thus, using the fact that $r_j = (1/p_j) - 1$, we can write the equilibrium odds, given that everyone acts only on his private information, $P_i(\omega)$, as

$$r_j = \varphi_j(\omega) = h \circ q^j(\omega)$$

where $h: \mathbb{Z}^n \to \mathbb{R}$ is defined by $h(z) = (1/\sum_{i=1}^{n} \beta_i z_i) - 1$ where $\beta_i = \alpha_i/(\sum_{k=1}^{n} \alpha_k)$. Clearly, h can be written in the form $h = g \circ f$, where $g(t) = (1/t) - 1$, and $f(z) = \sum_{i=1}^{n} \beta_i z_i$. But g is monotonic, and f is linear and monotonic, hence stochastically regular, so by Proposition 1, φ_j is of the form required for our theorem.

Example 4 Markets with Incomplete Information

This example considers the case of an economy with incomplete information, in which there is a complete set of state contingent futures markets, as developed in Grossman [1981]. Individual agents, each with private information, trade in futures contracts and then observe the market clearing price. So $S = \{s_1, \ldots, s_J\}$ is a finite set of J states of the world, and $A_j = \{\omega = (s,y) \in \Omega | s = s_j\}$, with $q_i^j(\omega) = \rho(A_j | P_i(\omega))$. We let $X = (\mathbb{R}^+)^{J+1}$ be a feasible consumption set, with typical element denoted $x_i = (x_{i0}, x_{i1}, \ldots, x_{iJ})$ and assume each individual has a state contingent utility function $u_i: X \times S \Rightarrow \mathbb{R}$, so that $u_i(x_i, s)$ represents i's utility for state contingent consumption of x_{i0} in the present, and x_{ij} in state j, given that state s occurs. Each individual also has a state contingent endowment $e_i \in X$ and a vector of shares $\theta_i = (\theta_{i1}, \ldots, \theta_{iF})$

of ownership in each of F firms. Now, once ω is drawn, each individual can compute posterior probabilities, conditional on his private information, of each future state of the world:

$$q_i^j(\omega) = \rho(A_j | P_i(\omega)).$$

Thus, given any price vector, $p = (p_1, p_2, \ldots, p_J)$, each individual can choose a consumption bundle $x_i \in X$ to maximize

$$E(u_i(x_i, s) | P_i(\omega)) = \sum_{j=1}^{J} u_i(x_i, s_j) q_i^j(\omega)$$

subject to the budget constraint imposed by the value of his endowment. Similarly, firms maximize profits at the given prices. (See Grossman [1981, p. 547] for details), resulting in the usual Walrasian market clearing price. We write $\varphi(\omega)$, for the Walrasian equilibrium price vector as a function of ω, where individuals act only on their private information, and do not try and learn from the price.

Under the special case, when utility functions are given by

$$u_i(x_i, s_j) = \alpha_i \ln x_{ij} + x_{i0}$$

and endowments satisfy

$$\sum_{i=1}^{n} e_{ij} = 1 \quad \text{for all } 1 \leq j \leq K,$$

Grossman [1981, p. 548] shows that the Walrasian correspondence satisfies, for each j

$$\varphi_j(\omega) = G(\sum_{i=1}^{h} \alpha_i q_i^j(\omega))$$

where G is a monotone increasing function. Clearly, we can write

$$\varphi_j = h \circ q^j$$

where $h: \mathbb{Z}^n \to \mathbb{R}$ is defined by $h(z) = G(\sum_{i=1}^{n} \alpha_i z_i)$ and $q^j: \Omega \to \mathbb{Z}^n$ is defined by $q_i^j(\omega) = \rho(A_j | P_i(\omega))$. Further we know from Example 3.1 that h is stochastically regular, so φ_j is of the form required for our theorem.

Discussion

All four of the above examples have a common structure. In each example, individuals have private information $\underline{P} = (\underline{P}_1, \ldots, \underline{P}_n)$, act on this information, yielding $x(\omega)$ and, for any $\omega \in \Omega$, this gives rise to public information $\varphi(\omega) = \eta \circ x(\omega) = h \circ q(\omega)$, which is exactly of the form required by our theorems. Therefore, we can apply our theorems to each example.

First, from Theorem 1, it follows, for all of the above examples, that when the public information, φ, is common knowledge, there is consensus on the underlying posterior probabilities. In examples 1 and 2, this means that $q_i(\omega) = q_k(\omega) = \rho(A | \hat{P}(\omega))$ for all $i, k \in N$. So in Example 1, when the aggregate information is common knowledge, all experts agree on the posterior probability of the event occurring, and in Example 2, when price is common knowledge, every firm agrees on the probability the chemical is toxic. In Examples 3 and 4, $\varphi(\omega) = (\varphi_1(\omega), \ldots, \varphi_J(\omega))$ is vector valued, so if φ is common knowledge at ω, then $\varphi_j(\omega)$ is common knowledge at ω for each j. Hence $q_i^j(\omega) = q_k^j(\omega) = \rho(A_j | \hat{P}(\omega))$ for all $i, k \in N$, and $1 \leq j \leq J$. Hence, in

xample 3, once the odds are common knowledge, then for <u>every</u> horse, ll bettors agree on the posterior probability that that horse will in. In Example 4, once the state contingent prices are common nowledge, then for <u>every</u> state, all agents agree on the posterior robability of that state occurring.

Second, from Theorem 2, if information partitions are finite, t follows that for all of the above examples, the iterative process efined in (5.2) converges to a common knowledge equilibrium haracterized by consensus.

To implement the iterative process of (5.2) for the examples bove, we would repeatedly, <u>without</u> <u>redrawing</u> $\underline{\omega}$, face the individuals ith the decision task described in each example. Thus, in each xample, we would have a series of decision periods. In period t, ndividuals choose optimal actions $x^t(\omega) = (x_1^t(\omega), \ldots, x_n^t(\omega))$ based on heir <u>current</u> information partitions $\underline{P}^t$. The aggregate information $\varphi^t(\omega) = \eta \circ x^t(\omega)$ is then made public, and based on this, individuals efine their information partitions to obtain $\underline{P}^{t+1}$, as described in (5.2). From Theorem 2, it follows that in <u>all</u> of the above examples, there is a t for which φ^t is common knowledge at ω under $\underline{P}^t$, and that for this t, we have consensus: I.e., $\rho(A_j | P_i^t(\omega)) = \rho(A_j | P_k^t(\omega))$ $= \rho(A_j | \hat{P}^t(\omega))$ for all $1 \leq j \leq J$ and $i, k \in N$.

It should be emphasized that in implementing this iterative procedure, there is only <u>one</u> realization of ω, and all subsequent decisions, $x^t(\omega)$, are taken with respect to the information that is generated by this ω. In each period, the action $x^t(\omega)$ must actually be taken, in order to generate information for subsequent periods. However, the payoffs cannot actually be made until the entire process is completed, since the payoffs themselves reveal information about the true state, ω.

In Example 1, the iterative process described above is exactly the "Delphi" process, in which individual experts make reports, an aggregate function of their reports is made public, they make new reports, etc. Theorem 2 shows that this process leads eventually to consensus. In Example 2, the iterative process described above corresponds to a situation in which there are several production periods, firms are faced with proportional liability for their share of the total production in each period, but they do not learn of the actual state (whether the chemical is toxic) until some later date. Here, Theorem 2 would suggest that such a proportional liability scheme would result in firms converging to identical beliefs about toxicity. It is somewhat more difficult to imagine natural methods of implementing the above iterative process in Examples 3 and 4.

Finally, for each of the above examples, we can define a REE. Using the notation of Section 6, we write $q_i^j(\omega,\underline{R}) = \rho(A_j|R_i(\omega))$ to indicate the posterior probabilities of A_j given $\underline{R} = (\underline{R}_1,\ldots,\underline{R}_n)$. In each example we denote the public information as a function of the initial information partitions by $\varphi(\omega,\underline{R})$. It follows from the same analysis that is done for $\underline{P}$, that in each example, we can write $\varphi(\omega,\underline{R})$ in the form $\varphi_j(\omega,\underline{R}) = h \circ q^j(\omega,\underline{R})$, where, now, from Theorems 1, 3, and

Corollary 2, we get the following results:

(a) If, for all ω, $\phi(\omega) = \varphi(\omega,\underline{P})$ is common knowledge at ω under $\underline{P}$, then ϕ is a REE (Theorem 3b) and for almost all $\omega^* \in \Omega$, for all j, $q_i^j(\omega^*,\underline{P}) = q_k^j(\omega^*,\underline{P}) = \rho(A_j|\hat{\underline{P}}(\omega^*)))$ (Theorem 1). Further, with finite information partitions, such a REE can be found by the iterative process of Theorem 2.

(b) If $\phi(\omega) = \varphi(\omega,\underline{P}_\phi)$ is a REE, then ϕ is common knowledge under $\underline{P}_\phi$ (Theorem 3a), and is characterized by consensus (Corollary 2). I.e., for almost all $\omega^* \in \Omega$, $\rho(A_j|P_{\phi,i}(\omega^*)) = \rho(A_j|P_{\phi,k}(\omega^*))$ $= \rho(A_j|\hat{P}_\phi(\omega^*))$ for all $i,k \in N$, $1 \leq j \leq J$.

In addition, the result of Corollary 3 is relevant to Examples 3 and 4 (although the information structure assumed there would only be likely to occur in Example 4). So we also have, for these examples:

(c) If $\underline{A} = \{A_j\}_{j=1}^{J}$ is a refinement of $\tilde{\underline{P}}$ with $\rho(A_j) > 0$ for all j, then every REE is fully revealing (Corollary 3). Thus, with finite information partitions, there is a unique REE, (existence follows from Theorem 2).

Thus, for each of the above examples, all REE's are characterized by consensus, and every common knowledge equilibrium gives rise to a REE with consensus.

In Examples 3 and 4, we obtain conditions on the information structure guaranteeing that there is a unique REE, and guaranteeing that REE is fully revealing.

A "paradox" of a fully revealing REE is that it appears to lack incentives for research (each i does just as well by conditioning on the public information alone and "throwing his private information away"). However, as Dubey, Geanakoplos, Shubic [1982] points out, and as is clear from this paper, there are indeed incentives for research during the iterative process.

Summarizing the above results, we see that our theorems have non-obvious implications in each of the above examples.

In Example 1, our approach gives a theoretical foundation for Delphi processes. In the literature on this subject, it has been claimed that such processes yield consensus among experts and are effective means of pooling the information of experts. However, there has been no adequate theoretical work from which such conclusions have been derived. Our analysis gives a way of addressing such questions theoretically, and our theorems give support for the claims made for such processes.

Our model of oligopolists, in Example 2, is formally equivalent—in the case of linear demand and only two states of the world—to models studied by Novshek and Sonnenschein [1972], Clarke [1984] and others, who assume the source of uncertainty is in the intercept of the demand function, rather than in the cost function. That literature shows that there are no incentives for firms to share information unless they can cooperate in the choice of production levels. Our analysis is at odds with this, in that we show that in a REE, firms acting non-cooperatively <u>can not help</u> sharing information,

at least to the extent that they will end up agreeing with each other on their beliefs on the relevant posterior probabilities. Thus, when firms act on their private information, they end up sharing it with others.

In the parimutuel betting example of Example 3, our analysis shows that in a REE, bettors achieve consensus in beliefs, and hence all bets are fair bets. (However, since bettors are not risk neutral, they still prefer to bet). Empirical studies such as Hoerl and Fallin [1974] suggest that the final totalizer odds <u>are</u>, in fact, close to fair bets (except for the track take).

In the example of markets with incomplete information, Grossman [1982] proves the existence of a fully revealing REE. However, he does not address questions of uniqueness or try to characterize REE's other than the fully revealing REE. The theorems of this paper give some tools for doing so. In particular, when preferences are as specified in Example 4, we show that every REE must achieve consensus, so that all agents agree on the probability of each state of the world. We can also identify certain conditions on the initial information partitions (there are undoubtedly other such conditions) which guarantee that every REE is a fully revealing REE, and hence guarantee uniqueness of the REE. To our knowledge such questions have not previously been addressed in the literature on rational expectations equilibria.

8. Conclusion

Markets distribute information as well as goods. And just as markets tend to bring preferences into agreement, on the margin, as individuals equate their marginal rates of substitution, markets may also align beliefs, as individuals make inferences from publicly available information such as prices.

Using the concept of common knowledge, we investigate the process of information aggregation and the tendency of beliefs to converge towards consensus. In addition to market settings, the analysis is applied to other situations such as Delphi processes and parimutuel betting.

FOOTNOTES

1. The meet of two partitions is their finest common coarsening. The join of two partitions in their coarsest common refinement.

2. The statement "i knows E" is interpreted to mean that $P^i(\omega) \subseteq E$. One could use instead the weaker condition that i knows E iff $\rho(E|P^i(\omega)) = 1$. In this case, the corresponding definition of common knowledge would not bear the same relationship to the meet, and the analysis is considerably complicated. We do not investigate this alternative definition here.

3. Note that the definition of monotonicity here is what is usually called strict monotonicity.

4. To formalize the notion that the probability space and information partitions are themselves common knowledge would require us to embed the entire problem in a larger product space in which first individuals are informed of the relevant probability space and information partitions of (5.1), and then of the particular partition element of $\underline{P}_i$ which occurs. We do not attempt to formalize this here, and in that sense the discussions in this and the following paragraphs is only intuitive.

5. To see that $[v'A^{-1}]_j$ is positive, note that A^{-1} has elements

$$a^{ij} = \begin{cases} \dfrac{1}{\lambda_j} - \dfrac{b}{(1 + Kb)\lambda_i\lambda_j} & \text{if } i = j \\ -\dfrac{b}{(1 + Kb)\lambda_i\lambda_j} & \text{if } i \neq j \end{cases}$$

where $\lambda_k = b + c_k$, and $K = \sum_{k=1}^{n} \frac{1}{\lambda_k}$. Hence $[v'A^{-1}]_j = \sum_{i=1}^{n} a^{ij}$ $= 1/\lambda_j(1 + Kb) > 0$.

REFERENCES

Aumann, R. J. "Agreeing to Disagree." Annals of Statistics 4 (1976):1236-1239.

Breiman, L. Probability (Reading, Mass: Addison Wesley, 1968).

Clarke, R. N. "Collusion and the Incentives for Information Sharing." Bell Journal of Economics 15 (Spring 1984):383-394.

Dalkey, N. C. The Delphi Method: An Experimental Study of Group Opinion (1969) RAND Corporation, Santa Monica, CA.

De Finetti, B. "Logical Foundations and Measurement of Subjective Probability." Acta Psychologica 34: Subjective Probability G. de Zeeuw et al, eds. (1970):129-145.

Geanakoplos, J. D. and H. M. Polemarchakis. "We Can't Disagree Forever." Journal of Economic Theory 28 (1982):192-200.

Grossman, S. J. "An Introduction to the Theory of Rational Expectations under Asymmetric Information." Review of Economic Studies 48 (1981):541-559.

Grossman, S. and J. E. Stiglitz. "Information and Competitive Price Systems." American Economic Review 66 (May 1976):246-253.

Hoerl, A. and H. Fallin. "Reliability of Subjective Evaluations in a

High Incentive Situation." J. R. Statist Soc. A (1974) vol. 137, part 2.

Jordan, J. S. "A Dynamic Model of Expectations Equilibrium." Journal of Economic Theory 28 (December, 1982):235-254.

Kobayashi, T. "A Convergence Theorem on Rational Expectations Equilibrium with Price Information." Working Paper No. 79, Institute for Mathematical Studies in Social Sciences, Stanford University, 1977.

Milgrom, P. and N. Stokey. "Information, Trade, and Common Knowledge." Journal of Economic Theory 26 (1982):17-27.

Novshek, W. and H. Sonnenschein. "Fulfilled Expectations Cournot Duopoly with Information Aquisition and Release." Bell Journal of Economics 13 (Spring 1982):214-218.

Plott, C. R. and S. Sunder. "Rational Expectations and the Aggregation of Diverse Information in Laboratory Security Markets." Social Science Working Paper #463, California Institute of Technology, (1983).

Savage, L. J. "Elicitation of Personal Probabilities and Expectations." Journal of the American Statistical Association 66 (December 1971):783-801.